Environmental Health Perspectives
Vol. 94, pp. 143–146, 1991

When Is It Time to Get Married? or When Should the Assay User and the Assay Developer Collaborate?

by Shanna H. Swan* and Bill L. Lasley*

Hormone assays are being developed in the laboratory to detect specific molecular markers in nonclinical populations. Epidemiology is increasingly using these assays to improve the precision with which disease processes and exposures can be defined. This growing body of molecular epidemiology requires a high degree of cooperation between the assay developer and the assay user. We draw on our experience in using a sensitive hormone assay for the detection of early pregnancy via urinary human chorionic gonadotropin to illustrate these points. We conclude that this collaborative effort, in addition to making this study possible, has provided unexpected rewards.

"But those epidemiologists don't understand the physiologic basis of these studies," the endocrinologist tells his colleagues. "Those endocrinologists are lost in detail; they study a few people endlessly," the epidemiologist reports.

Given the wide gulf between these views, can a "marriage" between epidemiology and endocrinology succeed? If the answer is in the affirmative, under what conditions is such a union likely to be successful? What is the likely product of such a union? We attempt to answer these questions using our experience in a collaborative effort between endocrinology and epidemiology in the study of early pregnancy loss. In particular, we describe conditions under which an assay is likely to be useful to the epidemiologist as a biomarker in large populations.

Background

Hormone Assays

Prior to the discovery and application of radioimmunoassays in the early and mid-1970s, analytic techniques such as colorimetry, gas-liquid chromatography, competitive protein binding, and double isotope dilution methods were employed to measure hormone concentrations. These tests were either too laborious or insensitive to be practical for routine monitoring. These methods were, however, quite useful for measuring changes in urinary hormone excretion because of the large quantities of urine that could be easily collected and the higher concentration of hormones in urine compared to plasma or serum.

With the advent of the radioimmunoassay (for which Berson and Yallow were awarded the Nobel prize), assays became more sensitive, and it was considered more appropriate to measure the active compound in circulation rather than its metabolite in the excreta. This, then, became the accepted standard in the field and replaced the use of urinary assays for research purposes.

Subsequently (1978–1980), urinary and fecal steroid hormone metabolite monitoring methods were developed in the zoo setting to obtain biological samples without capture and restraint. These second-generation assays took advantage of the improved technology that had been developed for serum and plasma assays and applied it to urine and feces (*1*). More recently, the recognition of the advantages of collecting and analyzing urine in human populations has renewed interest in this approach (*2*). Unlike blood sampling, urine sampling, which is nonintrusive, can be accomplished in large populations with a high level of compliance, samples can be collected over a prolonged time period, and collection requires no special skills or instruments.

Until recently, the assays employed to measure hormones have been radioimmuno- or radiometric assays applied to small clinic populations (*3*). Continued developments in this field have led to an array of assays that eliminate the need for radioactive materials, and in some case, obviate the need for a laboratory (*4,5*). The combined result of the concentrating effect of the kidney, the lack of binding proteins in urine, and the water-soluble nature of the hormone metabolites allows virtually all assay formats to be applied to unprocessed urine. Although dipstick kits are available for only a few hormones, more will be available shortly (*6*). The end result will be a series of easy-to-use, sensitive, and specific assays for use in large epidemiologic populations.

Epidemiology

Until recently, when sensitive urinary assays became available, epidemiologists interested in subfecundity and early pregnancy loss were limited to such indirect measures as the number and timing of births or the study of spontaneous abortion using

*Reproductive Epidemiology Program, California Department of Health Services, Berkeley, CA.
†Institute of Toxicology and Environmental Health, University of California, Davis, Davis, CA.

timing of births or the study of spontaneous abortion using clinically observable end points. However, it has long been suspected that clinically recognized fetal loss represents only a fraction of total postconception loss, although previous estimates of the magnitude of this loss have varied widely (*7,8*). Moreover, pregnancy loss is a relatively crude indicator of reproductive dysfunction; hormone levels, length of the luteal phase, and ovulation, if measurable, might provide more direct indicators of female reproductive health.

Even the study of clinically recognized spontaneous abortion is difficult. The earlier in gestation that a pregnancy is recognized by a woman or her health-care provider, the greater the probability that a spontaneous abortion will be reported. This is further complicated by the fact that women at high risk of loss, whether on the basis of past history, age, or symptoms in a current pregnancy, are more likely to seek early prenatal care. Therefore, losses in these high-risk groups are more likely to be documented. This form of selection bias is likely to be present in all epidemiological studies of clinically recognized pregnancy, regardless of design.

In addition, there are problems specific to the particular design used. Case-control studies must rely on self-reported pregnancy loss, which may be over- or underreported. Medical record review will minimize overreporting, but underreporting, particularly if differential, may remain a problem. Furthermore, retrospective recall of exposures that are likely to be poorly recalled may be problematic. On the other hand, cohort studies of recognized pregnancies may underestimate spontaneous abortion rates, since many losses occur either before the subject is identified as a cohort member or between the time the subject is identified and the interview. For example, spontaneous abortion rates in the literature, based on retrospectively recalled, self-reported loss, are 10 to 15% (*9,10*). In contrast, a recent case-cohort study found a crude spontaneous abortion rate of only 7.5% (*11*).

There is a way out of this dilemma: conduct careful surveillance on a group of women "at risk of pregnancy" and follow them uniformly to determine pregnancy outcome. There are several levels at which this surveillance can be carried out. The least sensitive of these methods would be to distribute home pregnancy kits to study subjects after they miss a menstrual period. A somewhat more sensitive method has been used in which women tested their own urine using home kits for a fraction of each month, whether or not a menstrual period had been missed (*12*). Alternatively, one can detect pregnancies as early as the seventh postconception day by conducting ongoing surveillance through sensitive urinary hormone assays carried out by an endocrine laboratory. This is the option we have chosen.

This choice was made possible by the recent development of a relatively inexpensive urinary hormone assay (*13*). Not only does this assay satisfy the usual laboratory requisites of sensitivity and specificity, but it also satisfies the conditions discussed below for successful cooperation between the laboratory scientist and the epidemiologist. This assay has successfully been used in one epidemiologic study (*14*) and is currently employed in several ongoing studies.

Prerequisites for an Effective Collaboration between Laboratory Scientist and Epidemiologist

Conditions on the Assay

It is of primary importance that the assay detect an end point that is both of epidemiologic and clinical interest. That is, the end point must be sufficiently common to warrant epidemiologic study. A rare genetic defect, which might have clinical significance, would not be of great epidemiological interest. Conversely, the epidemiologist is unlikely to find an assay developed unless it is clinically useful.

Needs for sensitivity and specificity differ somewhat for the laboratory scientist and the epidemiologist. The epidemiologist wants an assay with low false-positive rate (high specificity), for example, to avoid falsely telling women that they are pregnant. On the other hand, the epidemiologist, knowing that it is never possible to detect all pregnancies, may be less concerned about sensitivity than the laboratory scientist. However, it is important to both members of the team that the assay be highly reliable.

The assay should be easily usable in multiple laboratories, and all reagents should be available for a long period of time. For example, when the rabbit that produced the antibody used by Canfield (*13*) died, the resulting shortage of antibody delayed the use of radioimmunoassays for studies of early pregnancy loss. Fortunately, a monoclonal antibody that has the desired properties is now available.

To be useful epidemiologically, an assay should use specimens or samples that are easily obtainable, such as urine. The collection method should require a minimum of skill and equipment. For example, assays on bone marrow or fat tissue are clearly not useful in large populations. In addition, the assay should not require large volumes of specimens. For example, the original human chorionic gonadotropin (hCG) assays required 40 times the amount of urine now used. Current assays require less than 5 mL of urine, so it is now practical to ask women to store samples in their home freezers.

The collection procedure should be such that it can be imposed on subjects for a long time period. Based on only a few samples, it is not possible to establish "baseline" measurements, or to assess intrasubject variability. When infrequent events like pregnancy are under study, even more measurements are required. Subjects in the Wilcox and Canfield study (*13,14*) were required to collect urine for 6 months or until pregnant. A similar protocol is being followed in current studies. In addition, it is important that the timing of specimen collection and storage not be too critical. Fortunately, urine collection for hCG and steroid assays can be done at home, and these assays are not sensitive to temperature changes or freeze-thaw cycles. In fact, future urine assays may use a sample adsorbed onto filter paper, which will not require cold storage at all.

Finally, assay costs should be low enough to make studies in large populations feasible.

Conditions on Researchers

The distribution of the end points assayed in large populations must be of interest to the laboratory scientist as well as to the epidemiologist. The study of large populations rather than

individuals requires thinking in probabilistic and statistical terms. This frame of reference may be unfamiliar to the laboratory-based scientist. Conversely, the epidemiologist must be willing to think about the clinical implications of laboratory results, which may be equally unfamiliar. Thus, both members of the team must be willing to invest appreciable amounts of time learning about the field of mutual collaboration, so that interpretation of assay findings will be biologically and epidemiologically meaningful.

The primary interest of the laboratory scientist may well be assay development. However, for this collaboration to work, a laboratory must be set up to perform large numbers of repeated measurements. This involves the development of techniques for the meticulous labeling, handling, and storage of large volumes of samples. For example, a current study being conducted by the California Department of Health Services in collaboration with the Endocrine Laboratory at the University of California at Davis (the Women's Reproductive Health Study) will require handling close to 100,000 urine samples.

This task is being facilitated by the development of computerized data collection methods that can handle the accumulation of large numbers of laboratory samples. For example, for our studies, staff from the epidemiology and endocrinology groups have worked together to develop a bar-code-readable labeling and tracking system to facilitate storage and retrieval of the large volume of samples anticipated. This is a new area for both members of the team; epidemiologists are used to handling large volumes of questionnaires but not usually laboratory samples.

The epidemiologist must be willing to adapt field methods and even questionnaires to the demands of the assay. For example, for the Women's Reproductive Health Study, the endocrinologist has specified volume and frequency of sample collection, amount of sample collected, and the protocol for storage and transport of samples.

Others in the field (both assay developers and assay users) must also be willing to collaborate in order to standardize both laboratory and field methods so that results from multiple studies are comparable.

Conclusion

Under these conditions, the union of the laboratory assay with the epidemiological study can be very fruitful, with benefits, both expected and unexpected, to the epidemiologist and the assay developer.

Epidemiologists turned to studies of early pregnancy loss in order to *a*) detect early losses, which may be etiologically different from later losses; *b*) increase study power (in any fixed sample size) by increasing the number of measurable pregnancy losses; *c*) minimize misclassification of these outcomes; and *d*) decrease selection bias. All of these aims are likely to be met in our joint effort. It is likely that the measurable rate of pregnancy loss will be about three times higher in these studies than that observable without the use of hormone assays (*14*). Clearly, pregnancy losses defined by assay of biological samples will have much greater precision than if studied only through medical records or self-report. Studies conducted using chemical assay of biological samples should not be subject to selection bias, since the outcomes can be determined blindly both for the subject and the researcher.

In addition, we hope to learn about new health end points that might not be studied otherwise. Since this collaboration has begun, it has become clear that reproductive parameters, such as luteal phase abnormalities, anovulation, and decreased hormone levels, can be studied in addition to early pregnancy loss. These additional parameters may turn out to be even more useful and could not have been studied in large populations in the absence of urinary steroid assays.

The assay developer turned to epidemiology in order to understand the profile (mean, range, variability) of the assay in general populations. This information is necessary in order that clinical data in selected populations be properly interpreted. Our recent collaborative efforts should provide this information.

Additionally, through collaboration with the epidemiologist, the assay developer learns to think about the population rather than just the individual. In fact, the "unit of observation" has changed for the endocrinologists working on the Women's Reproductive Health Study. While individual samples were the unit of observation in the past, the endocrinologists now view the entire menstrual cycle as the unit of observation.

For researchers from both disciplines, data resulting from such a collaborative study will be more easily interpretable; biological mechanisms involved in any observed association should be clearer to the epidemiologist, and clinical abnormalities should be more understandable to the endocrinologist in the context of background rates.

The process of collaboration which our work on the Women's Reproductive Health Study has required, even prior to completion of data collection, has altered our thinking in profound ways. For example, prior to this collaboration, both disciplines viewed conception as an "all or nothing" phenomenon. It was a standing joke that you could not be "a little bit pregnant." With increased understanding by the epidemiologist of the subtleties involved in classifying cycles as "conception" or "nonconception" and an awareness by the endocrinologist of the power of probability theory to deal with such uncertainty, we now speak about the probability of conception in a given cycle. In fact, we are now allowing for a third alternative in our model, that of being "a little bit pregnant."

The principles discussed could equally well apply to assays detecting other markers of disease (e.g., DNA damage or sperm abnormalities) or markers of exposure (e.g., saliva assays for cotinine, nicotine, and caffeine). This kind of collaborative effort is likely to be of particular benefit to environmental epidemiology. In the past, environmental studies have often been weak because of poor assessment of exposure and outcome, both of which are often self-reported. The use of biomarkers provide a means of sharpening the tools available to epidemiologists. At the same time, it provides a population-based context in which researchers in the laboratory can interpret their assays. We look forward to continued fruitful collaboration of benefit to both the assay developer and the assay user.

REFERENCES

1. Loskutoff, N. M., Ott, J. E., and Lasley, B. L. Strategies for assessing ovarian function in exotic species. J. Zoo Anim. Med. 14:963–969 (1983).
2. Lasley, B. L., Stabenfeldt, G. H., Overstreet, J. W., Hanson, F. W., Szekala, N. M., and Munro, C. L. Urinary hormone levels at the time of ovulation and implantation. Fertil. Steril. 43: 861–865 (1985).

3. Armstrong, E. G., Ehrich, P. H., Birken, S., Schlatterer, J. P., Siris, E., Hembree, W. C., and Canfield, R. E. Use of a highly sensitive and specific immunoradiometric assay for detection of human chorionic gonadotropin in urine of normal, nonpregnant, and pregnant individuals. Clin. Endocrinol. Metab. 59: 867–874 (1984).
4. Munro, C. L., and Lasley, B. L. Non-radiometric methods for immunoassay of steroid hormones. In: Non-radiometric Assays: Technology and Application in Polypeptide and Steroid Hormone Detection (F. Hazelton and B. Albertson, Eds.), Alan R. Liss, Inc., New York, 1989, pp. 289–329.
5. Munro, C. J., Stabenfeldt, G. H., Overstreet, J. W., Cragun, J. R., Addiego, L. A., and Lasley, B. L. Serum and urinary ovarian steroids assessed by RIA and EIA. Clin. Chem. (in press).
6. Munro, C. L., and Lasley, B. L. A non-instrumented enzyme immunoassay (NEIA) to monitor ovarian activity by the measurement of urinary estrone conjugates. In: Non-radiometric Assays: Technology and Applications in Peptide and Steroid Hormone Detection (F. Hazelton and B. Albertson, Eds.), Alan R. Liss, Inc., New York, 1988, pp. 289–329.
7. Miller, J. F., Williamson, E., Glue, J., Gordon, Y. B., Grudzinskas, J. G., and Sykes, A. Fetal loss after implantation; a prospective study. Lancet ii: 554–556 (1980).
8. Edmonds, D. K., Lindsay, D. S., Miller, J. F., Williamson, E., Wood, P. J. Early embryonic mortality in women. Fertil. Steril. 38: 447–453 (1982).
9. Warburton, D., and Fraser, F. C. Spontaneous abortion risk in man: data from reproductive histories collected in a medical genetics unit. Am. J. Hum. Genet. 16: 1–25 (1964).
10. Rothman, K. J. Fetal loss, twinning and birthweight after oral contraceptive use. N. Engl. J. Med. 297: 468–471 (1977).
11. Hertz-Picciotto, I., Swan, S. H., Neutra, R. R., and Samuels, S. K. Spontaneous abortions in relation to tap water: an application of methods from survival analysis to a pregnancy outcome follow-up study. Am. J. Epidemiol. 130: 79–93 (1989).
12. Sweeney, A. M., Meyer, M., Aarons, J. H., Mills, J. L., and LaPorte, R. E. Evaluation of methods for the prospective identification of early fetal losses in environmental epidemiology studies. Am. J. Epidemiol. 127: 843–850 (1988).
13. Canfield, R. E., O'Connor, J. F., Birken, S., Krichevsky, A., and Wilcox, A. J. Development of an assay for a biomarker of pregnancy and early fetal loss. Environ. Health Perspect. 74: 57–66 (1987).
14. Wilcox, A. J., Weinberg, C. R., O'Connor, J. F., Baird, D. D., Schlatterer, J. P., Canfield, R. E., Armstrong, E. G., and Nisula, B. C. Incidence of early loss of pregnancy. N. Engl. J. Med. 319: 189–194 (1988).